Molecular Screening of VAX1 Gene Polymorphisms Uncovered the Genetic Heterogeneity of Non-Syndromic Orofacial Cleft in Saudi Arabian Patients

Jafar Sabbagh, Heba ; Innes, Nicola; Edris Ahmed, Sherif ; Butali, Azeez; Abdulbaset Alnamnakani, Eman ; Rabah, Sari

Published in:
Genetic Testing and Molecular Biomarkers

DOI:
10.1089/gtmb.2018.0207

Publication date:
2019

Document Version
Peer reviewed version

Genetic Testing and Molecular Biomarkers

Molecular Screening of VAX1 Gene Polymorphisms Uncovered the Genetic Heterogeneity of Non-Syndromic Orofacial Cleft in Saudi Arabian Patients

Journal:	*Genetic Testing and Molecular Biomarkers*
Manuscript ID	GTMB-2018-0207.R1
Manuscript Type:	Original Articles
Date Submitted by the Author:	24-Oct-2018
Complete List of Authors:	Sabbagh, Heba; Assistant Professor in Pediatric Dentistry, King Abdulaziz University, Jeddah, Saudi Arabia Innes, Nicola; Professor of Paediatric Dentistry, University of Dundee Dental School, Dundee, Scotland, UK Edris, Sherif; King Adulaziz university, Genetic medicine, Princess Al-Jawhara Al Ibrahim Center of Excellence in Research of Hereditary Disorders; King Adulaziz university, Biological sciences Department; Ain Shams University, Genetics Department Butali, Azeez; Assistant Professor. Department of Oral Pathology, Radiology and Medicine/ Dows Institute for Dental Research, College of Dentistry. University of Iowa. Iowa City, US Alnamnakani, Eman; Consultant Orthodontist, Program Director in Saudi Orthodontic Board Prince Sultan Military Medical City, Head of Cleft Lip and Palate Unit, Riyadh Military Hospital Riyadh, Saudi Arabia. Rabah, Sari ; Plastic and Reconstructive Surgeon, Chairman, Surgery Department, King Abdullah bin Abdulaziz University Hospital; Princess Noura University, Assistant Professor, King Saud bin Abdulaziz University for Health Sciences Hamdan, Mustafa; Consultant Plastic Surgeon, Plastic Surgery Department, King Saud Medical City, Ministry of Health Alhamlan, Nasir; Director, Pediatric Dentistry and Orthodontic Residency Program. Assistant Professor, King Saud bin AbdulAziz University for Health Sciences Abdulhameed, Fatma; Consultant Pediatric Surgeon, Pediatric Surgery Department, Madina Maternity and Children's Hospital Hassan, Mona; Professor of Biostatistics, Department of Dental Public Health, Faculty of Dentistry King Abdulaziz Unicersity; Professor of Biostatistics, Department of Dental Public Health, Faculty of Department of Biostatistics, High Institute of Public Health, Alexandria University. Al Mahdi, Hadiah; King Adulaziz university, Lab Coordinator Princess Al-Jawhara Albarhim Center of Excellence in Research of Hereditary Disorders, King AbdulAziz University Alamoudi, Najlaa; Professor of Pediatric Dentistry, Pediatric Dentistry, Faculty of Dentistry, King Abdulaziz University Al-Aama, Jumanah; Professor, Department of Genetic Medicine, Faculty of Medicine and Princess Al Jawhara Center of Excellence in Research of Hereditary Disorders. King Abdulaziz University Alaki, Sumer; Associate professor & consultant pediatric dentist. Vice

	dean of the Faculty of Dentistry, King Abdulaziz University Mossey, Peter; Professor of Craniofacial Development, Associate Dean for Research, Director of WHO Collaborating Centre, University of Dundee Dental School
Keyword:	Mutation Detection, Polymerase Chain Reaction (PCR), Genes, Genetic Testing
Manuscript Keywords (Search Terms):	NSOFC, SALIVA, VAX1, SAUDI ARABIAN, REAL TIME PCR, TAQMAN ASSAY

Molecular Screening of *VAX1 Gene* Polymorphisms Uncovered the Genetic Heterogeneity of Non-Syndromic Orofacial Cleft in Saudi Arabian Patients.

Heba Jafar **Sabbagh** BDS, MSc, PhD[1], Nicola PT **Innes** PhD, BDS, BMSc[2], Sherif Edris **Ahmed** PhD[3], Azeez **Butali** DDS, Ph[4], Eman Abdulbaset **Alnamnakani** MS, SB-ORTHO, MOrthoRCSD[5], Sari M **Rabah** MBBS, FRCSC, MHA[6], Mustafa A **Hamdan** FRCSed[7], Nasir H **Alhamlan** FDS RCSEd, MPH, MSc (Ortho), MOrth RCSEd, FICD[8], Fatma Dawood **Abdulhameed** MRCS, JBPS[9], Mona Hassan Ahmed **Hassan** MBChB, MPH, PhD[10], Hadiah Bassam **Al Mahdi** MSc[11], Najlaa M. **Alamoudi** BDS, MSc, DSc[12], Jumana Y. **Al-Aama** SBP, MRCP, FCCMG[13], Sumer M. **Alaki** MS, MPH, DrPH[14], Peter A. **Mossey** BDS, PhD[15]

Assistant Professor in Pediatric Dentistry, King Abdulaziz University, Jeddah, Saudi Arabia
hjsabbagh@gmail.com

Professor of Paediatric Dentistry, University of Dundee Dental School,
Dundee, Scotland, UK
n.p.innes@dundee.ac.uk

Assistant Professor of Molecular Genetics, Department of Biology, Genomic and Biotechnology Section Faculty of Science, Princess Al-Jawhara Al-Brahim Centre of Excellence in Research of Hereditary Disorders (PACER-HD)
Faculty of Medicine, King Abdulaziz University
Jeddah, Saudi Arabia
sedris@aucegypt.edu

Assistant Professor. Department of Oral Pathology, Radiology and Medicine/
Dows Institute for Dental Research, College of Dentistry. University of Iowa.
Iowa City, US
azeez-butali@uiowa.edu

Consultant Orthodontist, Program Director in Saudi Orthodontic Board
Prince Sultan Military Medical City, Head of Cleft Lip and Palate Unit,
Riyadh Military Hospital
Riyadh, Saudi Arabia.
dreman72@hotmail.com

Plastic and Reconstructive Surgeon,
Chairman, Surgery Department,
King Abdullah bin Abdulaziz University Hospital, Princess Noura University,
Assistant Professor, King Saud bin Abdulaziz University for Health Sciences,
Riyadh, Saudi Arabia
s_rabah@hotmail.com

Consultant Plastic Surgeon, Plastic Surgery Department,
King Saud Medical City, Ministry of Health,
Riyadh, Saudi Arabia
manh_2020@hotmail.com

Director, Pediatric Dentistry and Orthodontic Residency Program. Assistant Professor, King Saud bin AbdulAziz University for Health Sciences,
Riyadh, Saudi Arabia
alhamlann@yahoo.com

Consultant Pediatric Surgeon, Pediatric Surgery Department, Madina Maternity and Children's Hospital,
Madina, Saudi Arabia
dr.fdad@yahoo.com

Professor of Biostatistics, Department of Dental Public Health,
Faculty of Dentistry King Abdulaziz Unicersity, Jeddah , Saudi Arabia
Department of Biostatistics, High Institute of Public Health, Alexandria University.
Alexandria, Egypt
monaha59@hotmail.com

Research Assistant and Lab Coordinator
Princess Al-Jawhara Albarhim Center of Excellence in Research of Hereditary Disorders,
King AbdulAziz University,
Jeddah , Saudi Arabia
genetic_2006@yahoo.com

Professor of Pediatric Dentistry, Pediatric Dentistry, Faculty of Dentistry, King Abdulaziz University,
Jeddah , Saudi Arabia
nalamoudi2011@gmail.com

Associate professor, Department of Genetic Medicine, Faculty of Medicine and Princess Al Jawhara Center of Excellence in Research of Hereditary Disorders. King Abdulaziz University,
Jeddah, Saudi Arabia
jalama@kau.edu.sa

Associate professor & consultant pediatric dentist. Vice dean of the Faculty of Dentistry, King Abdulaziz University
Jeddah , Saudi Arabia
salagy@kau.edu.sa

Professor of Craniofacial Development, Associate Dean for Research,
Director of WHO Collaborating Centre, University of Dundee Dental School, Dundee, Scotland, UK
p.a.mossey@dundee.ac.uk

Corresponding author:

Heba Jafar **Sabbagh**

Assistant Professor in Pediatric Dentistry,
Faculty of Dentistry- King Abdulaziz University

P.O. Box 80209 Jeddah 21589

Tel Office: 6401000 (23580)

Mobile: 0505668481
Email: hjsabbagh@gmail.com

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

Abstract:

Objective: Non-syndromic orofacial cleft (NSOFC), including cleft lip with/or without cleft palate (CL±P) and cleft palate (CP) are multifactorial developmental disorders with both genetic and environmental aetiological factors. This study investigated the association between (CL±P) and (CP), and two polymorphisms previously determined using GWAS, as well as, the association between consanguinity and (CL±P) and (CP).

Methods*:* DNA using saliva was extracted from 171 affected individuals and 189 control group (age, gender and location) infant-parental triads, recruited from eleven referral-hospitals in Saudi Arabia. Two polymorphisms, rs4752028 and rs7078160, located on *VAX1* gene were genotyped using real-time polymerase chain reaction (qPCR). A transmission disequilibrium test was carried out using Family Based Association Test and PLINK to measure the parents-of-origin effect.

Results*:* Significant differences were found between affected individuals versus the control group. In the case of rs4752028 risk allele in cleft, the phenotypes were: CL±P (fathers: OR:2.16(1.38 -3.4); mothers: OR:2.39(1.53 -3.71); and infants: OR:2.77(1.77 -4.34)); and CP (fathers: OR:2.24(1.15 -4.36); and infants: OR:2.43(1.25 -4.7). For CL±P and the rs7078160 risk allele, the phenotypes were: (fathers: OR:1.7(1.05 -2.86), mothers: OR:2.43(1.49-3.97); and infants: OR:2.34(1.44 -3.81)). In terms of consanguinity, we found significant association between consanguinity and the rs4752028 polymorphism minor allele among CL±P compared to controls (P= 0.001).

Conclusion*:* This is the first study to find a relationship between the two loci on 10q25 (rs4752028 and rs7078160) and NSOFC in a population with high consanguinity.

Key words: Cleft lip, cleft palate, ~~10q25~~VAX1, consanguinity, aetiology

Introduction:

The aetiology of nonsyndromic orofacial cleft (NSOFC) includes cleft lip with/or without cleft palate (CL±P) and isolated cleft palate (CP) is complex. A combination of risk factors contributes to its aetiology; genetics, environmental and gene-environmental interaction (Mossey *et al*., 2009). ~~Recently, two large genome~~ Genome wide association studies (GWAS) identified 10q25 as a risk locus for CL±P (Beaty *et al.,* 2010; Mangold *et al.,* 2010). *Ventral anterior homeobox 1 (VAX1)*, a gene that codes for a protein that plays a role in the regulation of the body's developmental and morphogenesis processes, was reported to be associated with infants (product of consanguinity) affected by multiple craniofacial defects (Slavotinek *et al.,* 2012). In Saudi Arabia, a meta-analysis conducted in 2014 revealed that consanguinity is a risk factor for NSOFC (Sabbagh *et al.,* 2014). Therefore, two single nucleotide polymorphisms (SNPs) (rs4752028 and rs7078160) were considered plausible candidates for investigation of NSOFC in a community with a high prevalence of parental consanguinity as Saudi Arabia (el-Hazmi *et al.,* 1995).

The aim of this case triad-control triad study was to investigate the association between infant-parental rs4752028 and rs7078160 SNPs polymorphisms and both CL±P and CP in a Saudi population. We also investigated their relationships with risk of NSOFC phenotypes in the presence of parental consanguinity.

Materials and method:

Recruitment of Clinical Subjects

This paper is part of a series of studies on the prevalence of NSOFC (Abdulhameed *et al.,* 2014) and the aetiology of CL±P and CP in Saudi Arabia. Participants were recruited from three main cities; Riyadh (the capital city), Jeddah (the second largest city in Saudi Arabia), and Madina (one of the main cities in Saudi Arabia). Cases were recruited from neonatal

units, the plastic surgery departments and/or orthodontic clinics; all cases were examined in the department of genetic medicine. Criteria for subject selection included: 18 months or younger infants recruited from participating hospitals between January 1st, 2010 and December 31st, 2011. The total case-control study sample included 171 case triads and 189 control triads. The age, gender, and recruiting hospitals were matched in both the cases and control groups. The controls were healthy non-cleft infants that were selected at random from the neonatal or vaccination units. The control group, including parents and infants, were not affected by clefting of the lip and/or palate.

Triads with missing information or those who failed to give saliva samples were excluded from the analysis. In addition; infants with syndromes in case or control groups were excluded from the study.~~, those over 18 months of age or controls with parents who had orofacial clefts, were excluded from the study.~~

Ethical approval for this study was granted by King Abdulaziz University Hospital (359-10), Ministry of Health (C/47/302/38430), the Military Hospitals Institutional Research Review Board (IRB) (429/2011), and King Fahad Medical City (10-079). A questionnaire using yes/no questions was distributed to participants to collect personal information and ~~,~~ consanguinity information the details shown in supplementary table S1 and S2~~, and environmental factors associated~~. Interviews were conducted with parents to understand the type of consanguinity.

Clinical Sampling

Saliva samples were collected from infants and parents in both groups (case and control). Oragene kits were used for both samples; from the parents we used (OG-500), however, for the infants we used (OG-575).A consent form for both groups was signed by one of the parents.

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

Genetic Analysis

DNA was extracted by using QIAamp DNA Mini Kit (Catalogue # 51306). Quality and quantity measurement were evaluated using Qubit® 2.0 Fluorometer. Amplification of the two polymorphism, rs4752028 and rs7078160, was done using 7500 FAST Real-Time PCR (Applied Biosystem, Int.) by TaqMan®Genotyping assay and TaqMan Genotyping master mix (Applied Biosystem, int.). Samples were analysed by TaqMan® Genotyper Software (Applied Biosystem, int.) for scatter plot analysis. Supplementary Table ~~S1~~ S3 shows the characteristics of the two polymorphisms.

Statistical analysis

Hardy-Weinberg Equilibrium (HWE) tests were carried out using an online program (www.dr-petrek.eu/documents/HWE.xls) ~~(http://www.oege.org/software/hwe-mr-calc.shtml) (Purcell *et al.*, 2007)~~ to look for indications of inbreeding, population stratification, and problems in genotyping. This was carried out using chi-squared goodness of fit test with P-values of 0.05 to compare differences between the observed and expected values of the included homozygous and heterozygous genotype frequencies (Wigginton *et al.,* 2005). A transmission disequilibrium test (TDT) was carried out using Family Based Association Test (FBAT), and PLINK which was also used to measure the parents of origin effect. Comparison of polymorphism frequencies among CL±P and CP cases compared to controls were analysed using Chi Square test.

In addition, to detect which of the three types of polymorphisms provided the significant relationship, and to acknowledge the burden on type-1 error rate the threshold for declaring

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

statistical significance based on Bonferroni correction was determined to be p = 0.00056 using SPSS version 16.0 (SPSS Inc., Chicago, IL, USA).

The degree of association between allele frequency with CL±P and CP and also with parental consanguinity, compared to controls, were estimated by measuring the odds ratio (OR) and 95% confidence intervals (95% CI) using an online program (http://www.quantpsy.org/chisq/chisq.htm). OR and 95% CI were also used to measure the degree of association between rs4752028 and rs7078160 SNPs polymorphisms variants and parental consanguinity among oral cleft infants compared to controls. In addition, multinomial logistic regression was carried out to measure the interaction between consanguinity and genotype variant among NSOFC compared to control.

Results:

Out of the 171 NSOFC case-parental triads, 10 cases could not be grouped to a cleft phenotype because of missing information, resulting in; 161 nsOFC (127 CL±P; and 34 CP) cases for our analysis. In addition, 16 fathers out of the 189 control parental triads did not provide a saliva sample. First cousin marriages accounted for 55/86 (64%) of the NSOFC and 60/92 (65.2%) of the controls, out of the total parental consanguinity in these triads.

The case and control parental homozygous and heterozygous polymorphism frequencies in rs4752028 were aligned to HWE except for paternal control (0.039) . However, there were significant differences between the observed and expected values for both parental cases and controls at rs7078160 locus with p< 0.~~001~~05. ~~except for NSOFC fathers (p=0.060).~~ See Supplementary Table (S3~~2~~).

The transmission disequilibrium test (TDT) for rs4752028 and rs7078160, using FBAT and PLINK tests (Tables 1 and 2). No statistically significant over-transmission of the minor

allele (C in rs4752028 and A in rs7078160) was found in CL±P or CP families. In addition, PLINK tests found no parents of origin relationship (Supplementary Table S5~~3~~). For CP, the number of heterozygous alleles was insufficient to produce a P-value in FBAT analysis and the data were not included in Table ~~S3~~S5.

Comparison between case and control rs4752028 and rs7078160 genotypes and alleles

Table 3 shows the distribution of rs4752028 and rs7078160 genotypes in case and control infant-parental triads. There were statistically significant differences between cases and controls in (rs4752028 and rs7078160) genotypes in infant-parental triads for CL±P and CP cases.

After Chi Square adjustment using Bonferroni correction in infant-parent triads for rs4752028 SNP, in fathers; the homozygous TT common allele genotype was detected significantly more often in controls than in cases for CL±P and CP ($P< 0.05$). Furthermore, the heterozygous CT genotype was significantly more prevalent in cases than in controls for the different cleft phenotypes ($p< 0.05$). For mothers and infants, the homozygous CC minor allele genotype was significantly associated with CL±P cases compared to controls; the homozygous TT common allele genotype was detected significantly more often in controls compared to CL±P; and the heterozygous CT genotype was present significantly more often in CL±P and CP except in mothers of CP infants.

For rs7078160 SNP frequencies after Chi Square adjustment using Bonferroni correction, the homozygous AA minor allele genotype was significantly more frequent in CL±P infants compared to controls ($p<0.0056$). The heterozygous AG genotype was significantly more frequent in control infants compared to CL±P infants ($p< 0.0056$).

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

The frequency of the rs4752028 and rs7078160 minor alleles in case and control infant-parental triads CL±P and CP was compared. Significant differences between cases versus controls for rs4752028 risk allele in cleft phenotypes were: CL±P (fathers: OR:2.16 (1.38 - 3.4); mothers: OR:2.39 (1.53 - 3.71); and infants: OR:2.77 (1.77 - 4.34)); and CP (fathers: OR:2.24 (1.15 - 4.36); and infants: OR:2.43 (1.25 - 4.7). For CL±P and rs7078160, these were: fathers: OR: 1.7 (1.05 - 2.86), mothers: OR: 2.43 (1.49 - 3.97); and infants: OR: 2.34 (1.44 - 3.81). Consanguinity was significantly related to the rs4752028 polymorphism minor allele among CL±P compared to controls (P= 0.001, OR: 2.97 (1.54 - 5.76)). See supplementary ~~S4~~S6.

For rs7078160 SNP, there were statistically significant differences between CL±P cases and controls; fathers with significantly greater frequency of the minor A allele in CL±P cases compared to controls (P<0.05). However, this relationship was not statistically significant for CP.

Paternal consanguinity and infant rs4752028 and rs7078160 genotype variants as risk factors for CL±P and CP

CL±P, CP cases and controls were distributed according to parental consanguinity and then compared according to rs4752028 and rs7078160 infant-parental triad genotype variance. There were no statistically significant differences found in either analysis (p>0.05) (Supplementary Tables ~~S5~~ S7 and ~~S6~~S8). In addition, multinomial logistic regression with nsOFC as an outcome variable, consanguinity as main effect and phenotype as main effect with interaction term of the last two variables was carried out. It indicated significant main effect of genotype (P=0.0001 for rs4752028 and P=0.05 rs7078160) with no significant effect of consanguinity or interaction between them (P=0.5 for rs4752028 and P=0.2 for rs7078160)

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

Finally, when infants' rs4752028 and rs7078160 minor allele frequencies in CL±P and CP cases were compared to controls, there were more CL±P and CP cases with consanguineous parents and the minor C allele at rs4752028, but this was statistically significant for CL±P only (P= 0.001, OR: 2.97 (1.54 to 5.76). However, for rs7078160, although the minor A allele prevalence was higher in CL±P (13.4%) compared to controls (7.7%), the difference was not statistically significant (P=0.081, OR: 1.93 (0.92 to 4.04). See Table 4.

Discussion:

Our study showed statistically significant differences in the genotype variance and allele frequencies between CL±P and CP cases compared to control infant-parental triads. However, the FBAT and PLINK analysis did not show significant over-transmission of the rs4752028 and rs7078160 SNP alleles and parents of origin effect in NSOFC cases.

In this study, we selected cases and controls from the same hospitals, however, it was not possible to match ethnicity. Saudis, especially in the Western Region, have been of mixed ethnicity for hundreds of years. People from all over the world, of different ethnic origins, have travelled to Makkah and Madina on pilgrimage, then settled and mixed races through marriage. Additionally, Saudi Arabia has a unique geographic location between the three continents Asia, Africa and Europe, as a result of this, it can be difficult to group people according to their ethnicity in Saudi Arabia. Although they are generally considered Caucasian (Risch *et al.,* 2002). Moreover, Lewonin (2006) reported that every population has a separate geographic race and that they differ genetically to some degree from every other population. This emphasizes the need to carry out genetic research for each population.

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

Our sample met the HWE in rs4752028 SNP suggesting that our sample resembled expected population genotype frequencies for this study (Pritchard and Korf, 2008). However, both cases and controls at rs7078160 had significant differences between the observed and expected values of the included homozygous and heterozygous genotype frequencies suggesting the absence of random mating. This could be explained by the high prevalence of paternal consanguinity in the target population (el-Hazmi *et al*., 1995).

The infant case-control results do not represent independent replication of the results from parental groups. Moreover, due to consanguineous mating, the maternal and paternal case-controls results are not fully independent of each other. Therefore, this study examined and explained two methods of transmission of the alleles in question, the TDT and case-control analysis; and (TDT) remains robust for linkage in the presence of consanguineous populations. Autozygosity mapping might have been another consideration, which assumes the identical-by-descent co-transmission of mutations (Oliveira *et al.,* 2017), and this (also called consanguinity mapping) has not been applied to nonsyndromic OFC, in part because parental consanguinity is uncommon in places where research efforts have historically been carried out. This would assume no genetic heterogeneity and tight linkage of a disease gene with DNA markers.

An association between 10q25 locus and CL±P was supported by Leslie *et al*., (2017) who reported in their Genome-wide meta-analyses of nonsyndromic orofacial clefts that SNPS on 10q25 approached genome wide significance in NSOFC and CL±P groups among Asians.

Our rs4752028 and rs7078160 SNPs polymorphisms association finding was further supported by the results of Butali *et al.*, (2013) (15) in their replication of GWAS signals on 651 case-parental triads (Asian (494 infant-parent triads) and European (157 infant-parent triads) populations). FBAT analysis revealed a statistically significant strong association in

the transmission of the rs7078160 SNP among the Asian population ($p< 0.001$) but found no significant association in the European population, similar to our findings. However, their comparison of cases with controls in the Asian population showed an increased frequency of the common G allele compared to controls, which differed from our findings. Such differences could indicate ethnic and geographic variation between the Saudi population, which are Caucasians, and the Asian population in the genetic aetiology of CL±P.

Although the CP sample was small, it was still interesting to study a possible link with the included variants that could give preliminary information for planning future research. Rs4752028 was the only SNP examined that showed association with CP compared to controls ($p= 0.015$ for father, $p=0.049$ for mothers and $p= 0.009$ for infants). However, as the sample of CP in this study is considered small (34), this finding could only suggest a trend of association. Furthermore, Butali *et al.*, (2013) reported no significant association between rs4752028 and CP. Also, Duan *et al.,* (2017) reported parents of origin effect and no association between rs7078160 and rs4752028 SNPs and CP. However, their finding was concluded from TDT (FBAT) analysis and not from a case-control design. As *10q25* is a recently discovered locus in terms of risk for CL±P and CP, studies that clarify the relationship between NSOFC and rs7078160 and rs4752028 polymorphisms are still required.

A systematic review of parental consanguinity and NSOFC revealed a significant association (Sabbagh et al., 2014). At the same time, *VAX1* mutation was previously reported to be associated with birth defects in a sample with consanguineous parental marriages (Slavotinek et al., 2012). The relationship between rs7078160 and rs4752028 and consanguinity in case compared to control infants was analysed. For both SNPs, the minor allele was found more

often in CL±P cases with consanguineous parents compared to controls (Table 5). However, it was only statistically significant for rs4752028 (p= 0.001, OR: 2.97 (1.54 to 5.76).

Conclusion:

This is the first study to describe the relationship between two SNPS, rs4752028 and rs7078160, and NSOFC in a population with a high rate of consanguinity. There is an apparent association between rs4752028 and rs7078160 SNPs polymorphisms and both CL±P and CP in the Saudi population, but larger samples are needed for confirmation and definitive evidence. In addition, further investigation in the Saudi population, as well as, other populations is required to ensure consistency, and confirm the limits of the association study. Furthermore, it is not possible to expand this study to include other variants in or near the 10q25 loci or other genes, but due to resources limitations, it will be postponed. Therefore, future genome wide study, gene-gene interaction and gene-environmental interaction / epigenetics research is recommended to further clarify the aetiology of CL±P and CP. Confirmation of a positive association between consanguinity, NSOFC, and genetics will have a great implication for parental counselling and public health.

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

Declaration section:

Ethical approval for this study was granted by the King Abdulaziz University Hospital (359-10), Ministry of Health (C/47/302/38430), the Military Hospitals Institutional Research Review Board (IRB) (429/2011) and the King Fahad Medical City (10-079).

- The author(s) declare that they have no competing interests'.

- This project was funded by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah under grant No. (4/165/1431).

Acknowledgments:

This project was funded by the Deanship of Scientific Research, King Abdulaziz University, Jeddah (Grant No. 4/165/1431); and was conducted with the support of the University of Dundee Dental School World Health Organization Collaborating Centre for Oral Health and Craniofacial Anomalies. The authors also acknowledge the support of Princess Al-Jawhara Albrahim of Excellence for Hereditary Disorders (PACER-HD) where the laboratory experiments of this project were carried out.

The authors thank the research committees of the Ministry of Health in Riyadh, Jeddah, and Madina, the research committees and of King Saud Medical City, Riyadh National Guard Hospital, King Fahad Medical City, King Fahad Armed Hospital, and King Abdulaziz Medical City, and Dr Hassan Al-Naeem at King Fahad Hospital; Zamzam Ebrahim Al-Hakami and Nouf Al-Beshri at Al-Messadia Maternity Hospital; Dr Safinaz Salamah and Ebtisam Hussain at Al-Azizia Maternity Hospital; Mervat Ali Sayed and all nurses at King Abdulaziz University Hospital; Dr Mosleh Saad Alharbi, Dania Baeasa, and Dr Mamoon Daghestani at King Abdulaziz Medical City; and Dr Manal Al-Malik, Dr Fawzia Sabbagh, Dr Mawahib Abuauf, and Mariam Malope at King Fahad Armed Forces Hospital. Dr. Wamda Helal, Dr. Ahmed Mustafa Hamdan, and Dr. Bassem Mohamad Gesrha at King Saud Medical City.

References:

Abdulhameed FD, Sabbagh HJ, Hummaida TI, *et al.* (2014). Epidemiology of non-syndromic orofacial cleft (NSOFC) in Medina, Saudi Arabia. Exp Clin Cardiol 20: 505-516.

Beaty TH, Murray JC, Marazita ML, *et al.* (2010). A genome-wide association study of cleft lip with and without cleft palate identifies risk variants near MAFB and ABCA4. Nat Genet 42(6): 525-9.

Butali A, Suzuki S, Cooper ME, *et al.* (2013). Replication of genome wide association identified candidate genes confirm the role of common and rare variants in PAX7 and VAX1 in the etiology of nonsyndromic CL(P). Am J Med Genet A 161A(5): 965-72.

Duan SJ, Huang N, Zhang BH, *et al.* (2017). New insights from GWAS for the cleft palate among han Chinese population. Med Oral Patol Oral Cir Bucal 22(2): e219-e227.

el-Hazmi MA, al-Swailem AR, Warsy AS, *et al.* (1995). Consanguinity among the Saudi Arabian population. J Med Genet 32(8): 623-6.

Leslie EJ, Carlson JC, Shaffer JR, *et al.* (2017). Genome-wide meta-analyses of nonsyndromic orofacial clefts identify novel associations between FOXE1 and all orofacial clefts, and TP63 and cleft lip with or without cleft palate. Hum Genet 136(3): 275-286.

Lewonin RC. (2006). Confusion about human races: Race and genomics, social sciences research council. Available at http://raceandgenomics.ssrc.org/Lewontin/, accessed May 23, 2015,

Mangold E, Ludwig KU, Birnbaum S, *et al.* (2010). Genome-wide association study identifies two susceptibility loci for nonsyndromic cleft lip with or without cleft palate. Nat Genet 42(1): 24-6.

Mossey PA, Little J, Munger RG, *et al.* (2009). Cleft lip and palate. Lancet 374(9703): 1773-85.

Oliveira J, Pereira R, Santos R, *et al.* (2017). Homozygosity Mapping using Whole-Exome Sequencing: A Valuable Approach for Pathogenic Variant Identification in Genetic Diseases. The 10th International Joint Conference on Biomedical Engineering Systems and Technologies. Porto, Portugal, BIOINFORMATICS. 3: 210-216.

Pritchard DJ and Korf BR (2008). Medical genetics at a glance. Malden, MA, Oxford: Blackwell Publishing.

Purcell S, Neale B, Todd-Brown K, *et al.* (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81(3): 559-75.

Risch N, Burchard E, Ziv E, *et al.* (2002). Categorization of humans in biomedical research: genes, race and disease. Genome Biol 3(7): comment2007.

Sabbagh HJ, Hassan MH, Innes NP, *et al.* (2014). Parental consanguinity and nonsyndromic orofacial clefts in children: a systematic review and meta-analyses. Cleft Palate Craniofac J 51(5): 501-13.

Slavotinek AM, Chao R, Vacik T, *et al.* (2012). VAX1 mutation associated with microphthalmia, corpus callosum agenesis, and orofacial clefting: the first description of a VAX1 phenotype in humans. Hum Mutat 33(2): 364-8.

Wigginton JE, Cutler DJ and Abecasis GR (2005). A note on exact tests of Hardy-Weinberg equilibrium. Am J Hum Genet 76(5): 887-93.

Table 1 Transmission Disequilibrium Test (TDT) results for rs4752028 and rs7078160 variants among nsOFC infant parental triads and its phenotypes (CL±Pand CP) using Family Based Association Test (FBAT) analysis.

Type of nsOFC	Allele	afreq	fam#	P-value	OR and 95% CI
rs4752028					
CL±P	C	0.233	53	0.651	1.1 (0.71 to 1.71)
CP	C	0.221	14	1.00	1 (0.4 to 2.3)
rs7078160					
CL±P	A	0.128	45	0.327	0.76 (0.44 to 1.32)

afreq: Estimating allele frequencies

The transmission disequilibrium test (TDT) for rs4752028 and rs7078160, using FBAT and PLINK tests ***Table 1***

Table 1 : Testing rs4752028 and rs7078160 for transmission disequilibrium using PLINK analysis for nsOFC infant-parental triads and cleft phenotypes (CL±Pand CP).

nsOFC	Transmitted/ Untransmitted minor alleles	P-value	OR	A:U_PAR	P-value	Combined statistics P-value
rs4752028						
CL±P	41/37	0.651	1.11	01:01	1	0.655
CP	11/11	1	1	00:00	NA	1
rs7078160						
CL±P	22/29	0.327	0.759	02:01	0.564	0.414
CP	2/5	0.257	0.4	00:00	NA	0.257

A:U_PAR: Parental discordance counts by counting the number of alleles in affected versus unaffected parents
OR: Odd ratio
CL±P: Cleft lip with or without cleft palate
CP: Cleft palate

transmission disequilibrium test (TDT) for rs4752028 and rs7078160, using FBAT and PLINK tests **Table *1***).

Table 1 Distribution of rs4752028 and rs7078160 infant-parental triad genotypes according to nsOFC phenotypes (CL±Pand CP) and compared to controls.

Genotype	CL±P	CP	Control
rs4752028[a]			
Paternal genotype (frequency (%))			
Total	122	33	168
TT*	73 (59.8)	18 (54.5)	134 (79.8)
CT	44 (36.1)	15 (45.5)	29 (17.3)
CC	5 (4.1)	0	5 (2.9)
P-value	0.001**	0.001**	
Maternal genotype (frequency (%))			
Total	126	34	187
TT*	80 (63.5)	23 (67.6)	151 (80.7)
CT	36 (28.6)	9 (26.5)	32 (17.1)
CC	10 (7.9)	2 (5.9)	4 (2.1)
P-value	0.001**	0.180	
Infant genotype (frequency (%))			
Total	120	35	188
TT*	72 (60)	21 (60)	153 (81.4)
CT	39 (32.5)	13 (37.1)	32 (17)
CC	9 (7.5)	1 (2.9)	3 (1.6)
P-value	<0.001**	0.020**	
rs7078160[b]			
Paternal genotype (frequency (%))			
Total	119	34	167
GG*	86 (72.3)	26 (76.5)	141 (84.4)
AG	28 (23.5)	5 (14.7)	19 (11.4)
AA	5 (4.2)	3 (8.8)	7 (4.2)
P-value	0.150	0.320	
Maternal genotype (frequency (%))			
GG*	90 (71.4)	29 (85.3)	164 (86.8)
AG	27 (21.4)	3 (8.8)	19 (10.1)
AA	9 (7.1)	2 (5.9)	6 (3.2)
Total	126	34	189
P-value	0.004**	0.730	
Infant genotype (frequency (%))			
Total	122	35	186

GG*	90 (73.8)	32 (91.4)	157 (84.4)
AG	20 (16.4)	3 (8.6)	26 (14)
AA	12 (9.8)	0	3 (1.6)
P-value	0.003**	0.490	

* The homozygous common allele genotype
**The P value is significant at the 0.05 level.

[a] Eleven (6 cases and 5 controls) paternal samples, 5 (3 cases and two controls) maternal samples and 7 infant samples (6 cases and one control) did not produce genotyping values for rs4752028. The phenotype diagnosis for ten nsOFC cases are missing
[b] Fourteen (8 cases and 6 controls) paternal samples, one maternal sample and 7 infant samples (5 cases and 3 controls) did not produce genotyping values for rs7078160. The phenotype diagnoses for ten nsOFC cases are missing,
CL±P: Cleft lip with or without cleft palate, CP: Cleft palate

Table *1* shows the distribution of rs4752028 and rs7078160 genotypes in case and control infant-parental triads. There were statistically significant differences between cases and controls in (rs4752028 and rs7078160) genotypes in infant-parental triads for CL±P and CP cases.

Table 1 Distribution of infant rs4752028 alleles in cases and controls with consanguineous parents.

Allele type	CL±P	CP	Control
rs4752028 (*X*2=28.28, df=2, P<0.0001**)			
Total	130	43	182
T*	101 (77.7)	34 (81.4)	166 (91.2)
C	29 (22.3)	8 (18.6)	16 (8.8)
P-value	0.001**	0.059	
OR (CI)	2.97 (1.54, 5.76)	2.44 (0.97, 6.16)	
rs7078160 (X^2=6.11, df=2, P=0.047**)			
Total	134	42	182
G*	112 (86.6)	41 (97.7)	168 (92.3)
A	18 (13.4)	1 (2.3)	14 (7.7)
P-value	0.081	0.290	
OR (CI)	1.93 (0.92, 4.04)	0.29 (0.04, 2.29)	

**The P value is significant at the 0.05 level
* Common allele
CL±P: Cleft lip with or without cleft palate
CP: Cleft palate
OR (CI): Odd ratio and 95% Confidence interval

See **Table *1***

Table S1: Demographic characteristics of included sample

Demographic variable	NSOFC/ N=171	Control / N=189	P value
Location:			
Riyadh	62 (36.2%)	68 (36%)	0.83
Jeddah	71 (41.5%)	82 (43.4)%	
Maddina	38 (22.2%)	39 (20.6%)	
Child gender:			
Male	105 (61.4%)	114 (60.3)	0.441
Female	66 (38.6%)	75 (39.7)	
Age:			
Child	Mean: 6.39months SD:5.485	Mean: 7.99 SD: 5.542	0.008*
Father	Mean: 35.83 years SD: 8.545	Mean 34.79 years SD: 6.870	0.211
Mother	Mean: 29.38 years SD: 6.02	Mean: 28.89 years SD: 5.6	0.442

N: total number of families

n: total number of children

Significant at 0.05

Table S2: Distribution of the sample according to location, parental age and consanguinity

Variable	NSOFC/ N=171	Control / N=189	P value
Sample location:			
Riyadh	62 (36.2%)	68 (36%)	0.83
Jeddah	71 (41.5%)	82 (43.4)%	
Maddina	38 (22.2%)	39 (20.6%)	
Consanguinity			
Yes	86 (50.3%)	92 (49.5%)	0.76
1st cousins	55 (32.2%)	60 (31.7%)	0.93
Mean parental Age:			
Father	Mean: 35.83 years SD: 8.545	Mean 34.79 years SD: 6.870	0.211
Mother	Mean: 29.38 years SD: 6.02	Mean: 28.89 years SD: 5.6	0.442

N: total number of families

SD: Stander of deviation

Table S3. Polymorphism characteristics investigated

Gene Symbol	SNP ID	Chromosomal position	Variation	TaqMan Assay ID	[VIC/FAM]
VAX1	rs4752028	10:117075480	T>C	C__27883342_10	[C/T]
	rs7078160	10:117068049	G>A	C__31975118_10	[A/G]

Table S4. The observed frequency (OF) and expected frequency (EF) for rs4752028 and rs7078160 genotypes using the Hardy-Weinberg frequency calculation.

Groups	**Cases**			**Controls**		
Genotype	Common homozygous	Heterozygous	Rare homozygous	Common homozygous	Heterozygous	Rare homozygous
Paternal rs4752028 cases=165, controls=168						
Observed	95	65	5	134	29	5
Expected	98.5	58	8.5	131.3	34.47	2.26
X^2, P value	2.44, 0.12			4.23, 0.039**		
Maternal rs4752028 cases=168, controls=187						
Observed	104	52	12	151	32	4
Expected	100	58.8	8.6	149	35.7	2.14
X^2, P value	2.25, 0.133			2.03, 0.154		
Paternal rs7078160 cases=163, controls=165						
Observed	120	35	8	141	18	6
Expected	116	43	4	135.6	29.7	1.63
X^2, P value	5.666, 0.017**			19.07, 0.000013**		
Maternal rs7078160 cases=170, controls=189						
Observed	127	31	12	164	19	6
Expected	119.45	46.1	4.45	159.27	28.46	1.27
X^2(df), P value	18.24, 0.000019**			20.9, 0.000005**		

**The Chi-square statistic was considered significant at the 0.05 level

Table S5. Testing rs4752028 and rs7078160 for parent of origin using PLINK analysis for CL/P infant-parental triads and -phenotypes (CL±Pand CP).

	T:U_PAT	Paternal P-value	T:U_MAT	Maternal P-value	POO Z test	POO P-value
rs4752028						
CL±P	22.5:20.5	0.760	18.5:16.5	0.735	-0.047	0.962
CP	07:07	1	04:04	1	0	1
rs7078160						
CL±P	10:15	0.317	12:14	0.695	-0.443	0.658
CP	1.5:2.5	0.617	0.5:2.5	1	0.59	0.555

POO: Parents of Origin
T:U PAT: Paternal transmitted: untransmitted counts
T:U MAT: Maternal transmitted: untransmitted counts
OR: Odd ratio
CL±P: Cleft lip with or without cleft palate
CP: Cleft palate

Table S6. Distribution of rs4752028 and rs7078160 infant-parental triad allele frequencies in nsOFC, CL±Pand CP cases compared to controls.

Allele type	CL±P	CP	Control
rs4752028			
Paternal allele (frequency (%))			
Total	244	66	336
T*	190 (77.9)	51 (77.3)	297 (87.9)
C	54 (22.1)	15 (22.7)	39 (2.9)
P-value	0.001**	0.015**	
OR (CI)	2.16 (1.38, 3.40)	2.24 (1.15, 4.36)	
Maternal allele (frequency (%))			
Total	252	68	374
T*	196 (77.7)	55 (80.8)	334 (89.3)
C	56 (22.2)	13 (19.1)	40 (10.7)
P-value	<0.001**	0.049**	
OR (CI)	2.39 (1.53, 3.71)	1.97 (0.99, 3.93)	
Infant allele (frequency (%))			
Total	240	70	376
T*	183 (68.8)	55 (78.6)	338 (89.9)
C	57 (31.2)	15 (21.4)	38 (10.1)
P-value	<0.001**	0.009**	
OR (CI)	2.77 (1.77, 4.34)	2.43 (1.25, 4.7)	
rs7078160			
Paternal allele (frequency (%))			
Total	238	68	334
G*	200 (84)	57 (83.8)	301 (90.1)
A	38 (16)	11(16.2)	33 (9.9)
P-value	0.030**	0.129	
OR (CI)	1.73 (1.05, 2.86)	1.76 (0.84, 3.68)	
Maternal allele (frequency (%))			
Total	252	68	378
G	207 (82.1)	61 (89.7)	347 (87.8)
A	45 (17.9)	7 (10.3)	31 (8.2)
P-value	<0.001**	0.569	
OR (CI)	2.43 (1.49, 3.97)	1.28 (0.54, 3.05)	
Infants allele (frequency (%))			
Total	244	70	372
G	200 (82)	67 (95.7)	340 (91.4)
A	44 (18)	3 (8.6)	32 (1.6)
P-value	<0.001**	0.230	
OR (CI)	2.34 (1.44, 3.81)	0.48 (0.14,1.6)	

* The homozygous common allele genotype
**The P value is significant at the 0.05 level.
CL±P: Cleft lip with or without cleft palate, CP: Cleft palate, OR (CI): Odd ratio and 95% Confidence interval

Table S7. Distribution of infant rs4752028 and rs7078160 genotypes in cases and controls according to parental consanguinity.

Consanguinity	**CL/P**			**CP**			**Control**		
rs4752028									
Total infants	118[C]			33[C]			167[C]		
Genotype	TT*	CC	CT	TT*	CC	CT	TT*	CC	CT
Total	70	9	39	20	1	12	138	3	26 %
Yes%	60	66.7	43.6	70	100	50	55.8	100	46.2
No %	40	33.3	56.4	30	0	50	54.2	0	55.8
OR (95%CI)		1.3 (0.31, 5.8	0.52 (0.23, 1.14		a	0.43 (0.10, 1.89)		a	0.68(0.3, 1.57
X^2 (df), P-value	4.88 (2), 0.087			1.89 (2), 0.390			3.31 (2), 0.190		
rs7078160									
Total infants	116[C]			33[C]			166[C]		
Genotype	GG*	AA	AG	GG*	AA	AG	GG*	AA	AG
Total N	87	11	18	30	0	3	140	2	24
Yes %	60.9	54.4	33.3	66.7	0	33.3	56.4	100	41.7
No %	39.1	45.5	66.7	33.3	0	66.7	43.6	0	58.3
OR (CI)		0.77 (0.22, 2.72)	0.32(0.11, 0.94)*		a	0.25 (0.02 to 3.1)		a	0.64 (0.26,1.59
X^2 (df), P-value	4.62 (2), 0.099			1.31 (1), 0.252			2.55 (2), 0.279		

* Homozygous common allele genotype
**Significant level at P≤ 0.05
a. Not possible to analyze because the groups contain zero values

[C] Among rs4752028: two CL/P, two CP and 22 controls did not have their genotyping and/or paternal consanguinity information completed. Among rs7078160 [C] 10 nsOFC, 3 CL/P, one CP, and one control did not have their genotype and/or their paternal consanguinity information completed.

d Among rs4752028: there were two undiagnosed phenotypes. Among rs7078160: there were four undiagnosed nsOFC cases

Table S8. Comparison between case and control infant rs4752028 genotypes and their relationship to parental consanguinity.

Consanguinity	TT*			CC			CT		
	CL/P	CP	Control	CL/P	CP	Control	CL/P	CP	Control
rs4752028									
Total N	70	22	138	9	1	3	39	12	26
Yes %	60	63.6	55.8	66.7	100	100	43.6	50	46.2
No %	40	36.4	54.2	33.3	0	0	56.4	50	55.8
OR (95% CI)	1.9 (0.66-2.13)	1.3 (0.5-3.5)		a	a		0.9 (0.33-2.44)	1.17 (0.3-4.6)	
X^2 (df), P-value	1.17 (3), 0.770		1.33 (1), 0.514			0.41 (2), 0.813			
r rs7078160									
Total N	87	30	138	11	0	2	18	3	24
Yes %	60.9	66.7	56.4	54.4	0	100	33.3	33.3	41.7
No %	39.1	33.3	43.6	45.5	0	0	66.7	66.7	58.3
OR (95% CI)	1.2 (0.7-2.08)	1.29 (0.5-2.84)		a	a		0.6 (0.16-2.18)	0.6 (0.04-7.63)	
X^2(df), P-value	1.24 (2), 0.537		1.48 (1), 0.224			0.66 (2), 0.719			

* Homozygous common allele genotype

a. Not possible to analyze because the groups contain zero values

**Significant level at P ≤0.05